

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

Frank Festa

Welcome to On the Ballot with Ballotpedia, where we connect people to politics by providing neutral, nonpartisan and reliable information on our government, how it works and where it's headed. I'm Frank Festa. Thanks for being with us!

Have you noticed more text messages lately from polling agencies or researchers asking you to weigh in on any given contemporary political debate or storyline? Or maybe Joe Biden, Donald Trump or one of their celebrity supporters have been filling your spam folder with emails begging for donations. When this happens to me, I always wonder how exactly they got my email or phone number. And by now, we're all conditioned to expect a product advertisement after we go online shopping and leave something in our car. Somehow our personal information spreads all throughout the internet. And now we're not only getting ads for that new pair of running shoes, we're also getting ads on social media for competing products, complimentary products like fancy socks, and a plug for a popular running podcast as well.

But what if political campaigns are doing more or less the same thing with your data?

Today, I'm joined by the academic Kaushal Kafle, who is a current doctoral student at William & Mary and will be joining the University of South Florida in the fall as an assistant professor in the Computer Science and Engineering Department. Kafle and his colleagues analyze the privacy practices of political campaigns through their research of over 2,000 campaign websites for the November 2020 general election, Senate races, House races, even for the presidency. Unlike for -profit organizations, campaigns are considered non -profits and are not as heavily scrutinized when it comes to privacy practices.

Kafle studies the first large-scale analysis of its kind, full of interesting findings with important implications for this November election cycle.

Kaushal, thanks for coming on the show!

Kaushal Kafle

Happy to be here. Congrats again on graduating.

Frank Festa

Can I call you doctor yet or no? Is that premature?

Kaushal Kafle

I think you can. I don't know about the paperwork and stuff. I'm a defense, my final defense in a couple of days.

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

[Frank Festa](#)

I hope that goes well. I mean, sure, by the time that folks hear this, it might be already done with. So congrats again. We feel lucky to have you.

And lots of exciting stuff for us to talk about here. Obviously your paper has been circulating quite a bit here at Ballotpedia. You know, we saw a bunch of the mentions and we were really excited to talk with you more about what you found and what you learned throughout your research. So why do you think that there is such a significant lack of regulation regarding the privacy practices of political campaigns when compared to commercial entities?

[Kaushal Kafle](#)

Just to make a finer point on that, there is no federal data privacy regulation for profits. In the US, there are data privacy regulations, all of those are state regulations. So California has one, a bunch of other states like Virginia has one, but there is no comprehensive federal data privacy regulation. And I think the US suffers from that a little bit when it comes to even nonprofits, because there is no national conversation about the baseline for data privacy that customers can expect.

So I think the US suffers a little bit from that. Just when contrasting between for -profits and non -profits, what I would say is people just do not think of non -profits as data operations, like entities that collect a lot of data. And I think that being disconnected is what helps companies fly under the radar, even when they are collecting huge amounts of private data of customers in the background. So that disconnect is what was surprising to me when I came into this research.

The other thing is when it comes to political data in particular, it's a unique kind of data because political data, especially in the current climate, is very sensitive. It can reveal your party affiliations to what your socioeconomic beliefs are and stuff, and people may want to protect that. So at least our takeaway from this paper was there has to be a way to separate out the rest of the nonprofit entities with political entities or political campaigns because they belong to this domain where they handle like more sensitive data than others.

[Frank Festa](#)

Yeah. And we'll get to more of those takeaways in just a moment or two, but I'm curious, was there a story in the news that inspired you to investigate this further? Maybe it was your work in other security areas or was it something else?

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

Kaushal Kafle

So I don't know if you remember this, but in back in 2019, I think there was big news about Cambridge Analytica. It was a data harvesting operation or some of the election campaigns in the US and their opaqueness behind or their misleading way of collecting data from Facebook. All of those culminated into a big scandal. We were motivated by the fact that people actually started caring about more privacy around their political data or how their data is used within a political context.

And I think a survey after that, after the scandal, mentioned 76 % of customers caring about their data being used in a political context. So that was the beginning of it. We conducted a similar study in for-profit domains, which are more regulated in the US through state laws and internationally through regulations like GDPR, which recently came out in 2019, 2020.

So we did that analysis and we found a bunch of security and privacy gaps even there in entities that are supposed to be regulated. So we thought, you know, what does, if for-profit domain can have privacy gaps, what does that look like for a non-profit domain which, you know, may not have any standard at all?

Frank Festa

Yeah, it's an interesting way to think about it too, right? If those that are supposed to be regulated and monitored have gaps already, What about, what does that mean for those that are not regulated or observed in the same way? I feel like that's a pretty concerning sign.

I'm really curious about the way in which you guys conducted this research and how you collected all the information. There's so much good stuff in this paper and it's pretty easy to understand. So if folks would like to check it out, there is a link in our show notes for you, but let's jump right into some of the most interesting findings. So things that stuck out to me just generally, and I obviously want to hear some of what was most interesting to you, of course, but, for me, I thought it was really surprising that many privacy policies do not clearly state whether or not user data will be shared or sold.

And I think that that stuck out to me because a lot of the time when you get on a website these days, it'll give you your preferences for cookies. It'll give you a couple of disclaimers, you know, tucked into whatever corner of your screen, whether it's a phone or a laptop or what have you. But yeah, I thought it was really surprising that the same is not being done for these types of campaigns. What are some of the potential consequences of this lack of transparency for users and what did you learn just by noticing that?

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

Kaushal Kafle

So there are a couple of issues there with regards to sharing. So when, again, election campaigns are unique in the fact that users may want to voluntarily give their information to some of the campaigns that they support. You can see that election campaigns are being entrusted with you like voters data.

in a very unique sense. And the fact of the matter is users may expect their data to be shared, but if there is no privacy disclosure being mentioned, that means that election campaigns are not even creating a standard or expectation behind how they can use that data. So if there is nothing written in Word, then they can sort of like use that data for whatever purpose that they want and it may go beyond what the voters are expecting.

Frank Festa

Kausal, if you don't mind, I'm sure there are some instances and references that came up to you in the research. What stuck out? Do you have an example for us?

Kaushal Kafle

Yeah. So I think there was a campaign in 2016 election, one of the presidential campaigns actually sold voter information, voter data to a rival campaign after the campaign was over. Yeah. So things like that.

may not be what the voters initially submitting the information would be signing up for. They may not expect that to happen with their data, even in cases where they support the campaign directly. So stuff like that, if there is no expectation of privacy through an agreement with the user, then it sort of gives the campaign leeway to do whatever they want with the data after the campaign ends. And on that point,

One of the most critical, I would say, finding of the paper was that none of the campaign, none of the, I think we analyzed more than 500 privacy policies in this study. And none of the campaign mentions what happens with the data after the campaign ends. We all know that campaigns are supposed to end after a period of time. And there was no mention of any retention policy after the campaign ends.

Do they sell the data? Do they sell the data? We don't know.

Frank Festa

Well, tell me why that's important too. Is there, and that's an interesting way to look at it, right? Like, let's say I'm running for, you know, Senate here in New York where I live. I lose, right? Stinks. I gotta go back to my day job. I'm going to shut the website down. You know, the campaign didn't work, but is there some sort of like website graveyard?

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

where there's some level of concern where this information is still technically out there on the internet, right? And still can be captured or sold or monetized in some way?

Kaushal Kafle

So data submitted to election campaign websites, they are stored in the campaign websites databases somewhere. So technically they can go back to that database and do whatever they want with the whatever campaigns can leverage that data for whatever purpose they want after the campaign ends.

So that is one way of looking at it. The other one is we hear so frequently about websites being hacked and databases being leaked. So once you put the data somewhere, we don't know what the security of that apparatus looks like. So motor's data is almost perpetually in risk after the point of submission. So because there is no transparency about what exactly is being done with that data. So there is almost like no way of knowing from the user side how long we should be on guard for that data going forward. And it is most likely going to be in perpetuity.

Frank Festa

Yeah. I just want to call out the percent as well that you shared with me before we hopped on our interview here just now. I believe it was 71 % of the campaigns that you analyzed had no privacy policy listed at all. Isn't that right?

And that's, that's a pretty substantial percentage. You know, I feel like you were probably really surprised when you saw that, right? Or is that kind of what you were expecting? Did you know that this is what you were getting into?

Kaushal Kafle

I mean, we kind of expected that, that it would be worse than the, the for -profit domains. And we had analyzed our for -profit domain previously to, to set our expectation a little bit, but it turned out to be way, way less percentage of the campaigns actually had data, had privacy policy compared to the for -profit domain of smart home vendors. So that was what was surprising, the degree.

Another stat that stuck out to me was that 41 % of campaign websites use third party trackers, which is something that I feel like we hear about a lot. But again, like if you asked me to define it on the spot right now, I'm not sure I would get that Jeopardy question right. So what does that mean? Why is that important? What are the implications for user privacy, like what are these third party trackers do?

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

Kaushal Kafle

Yeah. So, when we hear trackers, we may not immediately know, you know, we may be alarmed by just, just the fact that there are trackers in a website. Most of the trackers that are in a website are, are going to be what are called analytics for, for analytics. So where is the, the user looking at what, what, what items, what entities did the, in the website, did the user click on?

What is the IP address of the user that is viewing their website? So most of the time it is the metadata associated with the browsing habits or the browsing metadata associated with the device that the user is accessing the website from. So the reason why trackers are important in this context is it is another data point for a campaign on top of everything that they're collecting directly through their input forms.

And when that is combined with the fact that there is no transparency or there is no privacy disclosure in that website, we don't actually know what kind of information is being tracked from the trackers unless the campaigns specify that in the privacy disclosure. So again, the lack of transparency is the problematic part here.

Frank Festa

Yeah, and the lack of transparency and all that ambiguity really creates a very, very deep rabbit hole to imagine wondering.

What is this information being used for? Where's it going? And I'm curious why you were looking through all of these over 2000 campaigns. Did you find anything that could have, you know, concretely told you what's going on with that data? Like, did you find any instances where, you know, you found a record of some data being sold for XYZ purpose, maybe to a polling firm or somebody else or a marketing agency or anything like that?

Kaushal Kafle

There is one element of the research that we focused on was to actually find real -world insights into how campaigns operate. So what we did was we signed up our data to 26 campaigns, I believe. Some in the House candidates, some in the Senate candidates, and some in the presidential candidates' websites. So what we found was, I mean, our expectation was that our data was going to be shared, and we wanted some real -world insight into who that is if there is data sharing going on, who that data is going to be shared with. And the way we did this was we created 26 separate email addresses for each of the campaigns, completely new. So there is no, we didn't use that email address anywhere else. So we could monitor the inbox, any incoming email we would potentially get.

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

What we found was I think there were eight campaigns that we got emails from entities that we hadn't directly signed up for. And among them, I think three of them did not have a privacy policy or told us that our data was going to be shared. So even among the 26 that we directly signed up for, we already had some real -world evidence that our data was being shared without being disclosed to us.

Frank Festa

What else can you tell us about your methodology that might be interesting to our listeners? I know again, we brought up a few times, but looking at 2000 versions of anything is a Herculean task. So what did that process look like?

Kaushal Kafle

I do want to give a big shout out to Ballotpedia in this particular instance, because Ballotpedia was a big resource that we used to automate this whole process. To be able to analyze...anything like, as you said, like more than 2000 is a big task. And we wanted to go bigger. Like we wanted our approach to scale because of the fact that we wanted our study to be a general snapshot of what was happening. We didn't want a particular case study that we had to extrapolate. We wanted a clear snapshot of what was happening. So how we did this? How we developed this methodology was to combine the FEC database. So FEC, the Federal Election Commission, requires all the candidates to register their campaign with the federal government. That is where you can get information about all of the registered campaigns that will participate in the election. So that is where we started. But what we did was we combined all the metadata of campaigns that was there with Ballotpedia information. So the reason Ballotpedia was important was FEC requires obviously campaigns and candidates to use their official name. So Joe would be Joseph, for example. So it was a hard task to directly get the FEC name and then find the campaigns website because capital website generally has informal name of the candidate. So what we did was we, yeah, we docked to post the FSC data set with the Ballotpedia information that we find to first extract the, the Ballotpedia profile of campaigns. And once we have the Ballotpedia profile of campaigns, we automatically pass through the webpages to get, get the campaign URL. And then eventually we, we developed our tools to automatically download the campaign website after that.

Frank Festa

So, and I just want to be clear really quick. We did not pay Kosel to say that Ballotpedia helped him out here. That was a genuine plug and we really appreciate it. It's so exciting to see the data used in that way, you know, but I wanted to ask, you know, you're taking a look at the 2020 campaigns for the most part. It's 2024. How are you thinking about your findings and what you've learned in light of our living elections? What's your frame of mind?

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

Kaushal Kafle

The reason why I felt this study was important was to set light on the, what's the light on behalf of the voters to see what was going on. And we really hope that that campaigns, first of all, like voters have better expectation from campaigns, which is the only way to make an impact here, like to improve the status quo. When, when voters start having expectation from their campaigns,

that they support regarding data privacy, that is when we can expect improvement from the campaigns. The other thing to keep track of is there has been a lot of movements when it comes to data privacy in the EU. So I'm also looking out for campaigns that can make campaign pledges regarding data privacy regulation that they're looking to support in the future. Because I feel like the US does need almost like a federal level conversation regarding what, what data privacy should be or a normal consumer, what they can expect.

Frank Festa

And this might seem obvious to you, Kosal, but why should people care? I know it's the million dollar question for every research endeavor, or I know you're gearing up to, you know, teach a class full time down in South Florida. So better get ready to start explaining things and making that sales pitch over and over again. I'm curious, what's the argument for why everyday folks might care? If I'm visiting a political campaign website, what should I be aware of in terms of data privacy?

Kaushal Kafle

So regarding the question about why people should care, I think most of the data harvesting entities in the US benefit from the fact that there is no expectation of privacy from US consumers. And unless we want, there has been, and this is backed up by research prior work, that when users are actually shown how their data is being used within a system, that is when they start caring about, you know, there should be privacy disclosures, there should be more transparency. Because that is when they know that my data can be used in these many scenarios. So it is often that gap of thinking how your data is being used that needs to be brought to users' attention, and that is what makes them jump over the hoop of actually starting to care about their data. So that is one thing.

The other thing is, especially in the political context and especially in the current climate, there is a lot of changes that can happen after the user submits their data to a specific election campaign. Maybe the voters change their opinion about a campaign and they want their, they regret sending their data to a particular campaign or the campaign changes their affiliation down the line, given the climate right now. Or, you know,

OTB TRANSCRIPT – Episode #111

Personal data and the privacy gaps of political campaigns

just from the fact that none of the campaigns outline what happens to a voter's data after it is submitted, after the campaign ends. I think the users sort of start demanding, or at least setting an expectation, like sending your data to someone with an expectation that, okay, the use of this data is never going to be transparent to me. I think just knowing that expectation, keeping that expectation in your head before submitting your data would be the first point that we need to be aware of. Knowing the potential for issues down the line should be aware of. And weighing the benefit of submitting a data versus the risk of submitting that data. Once we have an informed user base, then we can expect a more informed voter base going forward so that we can help the users make informed choices regarding the campaigns that they interact with. their online data privacy habits and so on. So I feel like, you know, from the user's perspective, there's a lot to care about.

Frank Festa

No, definitely, and campaigns are in full swing again. We're just a few short months away from the election. So doubtlessly, all this is going on right now, right under our noses. So we'll have to stay tuned and see how everything develops. But Kaushal, thank you again, man. I really appreciate you coming on the show and for your insights. It was really interesting and I hope our listeners felt so too. So thank you again.

Kaushal Kafle

Yeah, thank you so much. I really enjoyed it. Thank you.

Frank Festa

And for our listeners, you can learn more about our election coverage at the links in our show notes. We'll be back next week with another episode. Make sure you subscribe to On The Ballot wherever you listen to podcasts. I'm Frank Festa and thanks for listening. We'll see you next week.